ABSTRACT

Visual feature information which is information representing a numerical value of a visual feature of an object and additional information which is information added to the object are stored in association with each other. Partial image data which is image data of a partial area of a video image is extracted. Visual feature information of the extracted partial image data is generated. The visual feature information of the extracted partial image data and visual feature information of an object which is stored are compared with each other to calculate a similarity therebetween. Based on the calculated similarity, an object contained in the video image data is identified. An annotation made up of additional information of the identified object is displayed in superposing relation to the video image on a display device.